



Aspect-Based Sentiment Classification Using Deep Learning and Hybrid of Word Embedding and Contextual Position

Waqas Ahmad¹, Hikmat Ullah Khan^{1,2,*}, Fawaz Khaled Alarfaj^{3,*}, Saqib Iqbal⁴,
Abdullah Mohammad Alomair³ and Naif Almusallam³

¹Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah, 47040, Pakistan

²Department of Computer Science, Namal University, Mianwali, 42201, Pakistan

³Department of Management Information Systems (MIS), College of Business Administration, King Faisal University (KFU), Al-Ahsa, 31982, Saudi Arabia

⁴Department of Software Engineering, Al Ain University, Al Ain, 64141, United Arab Emirates

*Corresponding Authors: Hikmat Ullah Khan. Email: hikmat.ullah@ciitwah.edu.pk;
Fawaz Khaled Alarfaj. Email: falarfaj@kfu.edu.sa

Received: 25 March 2023; Accepted: 06 June 2023; Published: 11 September 2023

Abstract: Aspect-based sentiment analysis aims to detect and classify the sentiment polarities as negative, positive, or neutral while associating them with their identified aspects from the corresponding context. In this regard, prior methodologies widely utilize either word embedding or tree-based representations. Meanwhile, the separate use of those deep features such as word embedding and tree-based dependencies has become a significant cause of information loss. Generally, word embedding preserves the syntactic and semantic relations between a couple of terms lying in a sentence. Besides, the tree-based structure conserves the grammatical and logical dependencies of context. In addition, the sentence-oriented word position describes a critical factor that influences the contextual information of a targeted sentence. Therefore, knowledge of the position-oriented information of words in a sentence has been considered significant. In this study, we propose to use word embedding, tree-based representation, and contextual position information in combination to evaluate whether their combination will improve the result's effectiveness or not. In the meantime, their joint utilization enhances the accurate identification and extraction of targeted aspect terms, which also influences their classification process. In this research paper, we propose a method named Attention Based Multi-Channel Convolutional Neural Network (Att-MC-CNN) that jointly utilizes these three deep features such as word embedding with tree-based structure and contextual position information. These three parameters deliver to Multi-Channel Convolutional Neural Network (MC-CNN) that identifies and extracts the potential terms and classifies their polarities. In addition, these terms have been further filtered with the attention mechanism, which determines the most significant words. The empirical analysis proves the proposed approach's effectiveness compared to existing techniques when evaluated on standard datasets. The experimental results represent our approach outperforms in the F1 measure with an overall



achievement of 94% in identifying aspects and 92% in the task of sentiment classification.

Keywords: Sentiment analysis; word embedding; aspect extraction; consistency tree; multichannel convolutional neural network; contextual position information

1 Introduction

During the last decades, a boom in the social media content explored on web-based platforms, which incorporated users' attitudes, reviews, and opinions towards various products or services. These feedbacks are advantageous for companies, organizations, and producers to elevate the quality of their products and services. In this scenario, Sentiment Analysis (SA) proved to be a powerful companion for companies, organizations, and producers. The reason is the ability to extract consumers' attitudes, reviews, and opinions towards their products and brands' related aspects [1,2]. The main objective of SA is to classify the web-based content comprised of text- sometimes audio- and video into neutral, negative, or positive. According to relevant literature, SA assumed that whole-textual content (document or sentence) expresses a consistent or unified polarity that does not hold in a real-world scenario. Aspect-Based Sentiment Analysis (ABSA) has overcome this drawback, which believes that sentiment polarity relates to aspects or features of any product or entity. Thus, this task accomplishes within two steps; the first step includes the extraction of features, while the second step includes the assignment of opinions corresponding to all the extracted attributes or aspects [3,4]. Aspect extraction and their sentiment analysis have been considered well-known subtasks of ABSA and are within the scope of the proposed approach. Generally, these aspects are of two types. These are implicit (not openly mentioned in textual reviews) and explicit (clearly defined in user reviews). As a result, the proposed approach considers the extraction of explicit aspects, whereas implicit aspects are not in the scope of this paper.

The traditional methodologies utilize varied approaches such as topic modeling-based, lexicon-based, machine learning-based, and rule-based or syntactic relation-based to accomplish the task of explicit aspect extraction and analyze their relevant sentiment. In addition, most of these methodologies have been discouraged due to their manual feature engineering mechanisms, time consumption constraints, demanding specialists for rules creation, and depending upon massive labeled datasets [5]. Nowadays, one of the most significant machine learning techniques is Deep Learning (DL), which comprises the ability to perform automatic feature selection without any manual intervention. Under these consequences, their algorithms express improved efficiency from preceding methods such as machine learning, linguistic rule/pattern, and lexicon oriented methods. However, such approaches are highly dependent in either sequential structure or tree-based representations to extract the potential aspects and relate them to their corresponding sentiments. Therefore, word2vec embedding utilizes to obtain the sequential illustrations because it can capture the semantic and syntactic information of the textual data. While on the other hand, tree-based representations such as the dependency relation-based tree structure and consistency tree structure are utilized to formulate the grammatical and logical relations of context. Regardless the grammatically dependent dependency tree performs well in aspects and their related sentiments identification. However, neither users nor websites care about grammatical rules when users express their opinions and comments on the web portals that accept their feedback on social media platforms. In that situation, the consistency tree is still performing well, even when

the grammatical rules are being violated, in identifying those categorical parts that originated these textual phrases [6,7].

Among the DL algorithms, Recurrent Neural Networks (RNN) and Convolutional Neural Network (CNN) are the two most well-reputed algorithms. They have proven their remarkable performance in the Natural Language Processing (NLP) scenarios. Under these consequences, these algorithms target the field of ABSA and accomplish the main tasks of aspect extraction and sentiment analysis [8,9]. However, RNN depicts incredible performance while extracting aspects and relevant sentiments compared to Conditional Random Field (CRF). While, on the other hand, some limitations were observed in the relevant literature regarding RNN performance when they handle tree-based dependencies of the sentence while accomplishing the main task of ABSA. The actual reason behind this is that RNN developed as a sequential model, which does not have sufficient abilities to handle such rich dependencies [10]. In the meantime, CNN has proven its outperformance while accomplishing the task of NLP, e.g., classification of text, matching words semantics, context-dependency, and preserving sequential order. Therefore, these tasks depict the adequate abilities of CNN that are enough to handle the tree-based sentence dependencies [11].

However, these algorithms have massively relied on either word2vec embedding (which contains the syntactic and semantic information) or tree-based structure (which includes the grammatical and logical relations information) for aspects' identification and classifying their sentiment [12]. In addition, only dependence upon the word2vec embedding becomes the cause of ignoring the grammatical and logical relations of the sentence. On the other side, dependence only on the tree-based structure becomes the cause of neglecting the syntactic and semantic information of the targeted sentence. Both these representations ignore the contextual position information, which expresses the valuable knowledge about the degree of influence concerning different word places in the context of a sentence. Therefore, these considerations can enhance the identification accuracy of features (aspects) and sentiment prediction within context, which still demands the attention of the research community [13,14]. Recently, researchers presented numerous techniques that address the aspect identification and classification task with collaboration of attention mechanisms. It enhances sentiment classification with credible opinion targets exposition. In the meantime, the exploration related to the interaction between aspects and contextual words with the accurate position of influenced terms can enhance the abilities of attention mechanisms [15,16].

As a consequence, from the above discussion, this study concludes that separate utilization of word2vec and tree-based structure causes information loss. That reason motivates this study to propose an approach that jointly utilizes word2vec embedding and a tree-based consistency structure for identifying potential aspects and their sentiment. The motivation behind this is that word2vec embedding provides both syntactic and semantic information about the sentence. Meanwhile, the consistency tree structure helps the model during the identification of those categorical elements, which formulates different sentences. Additionally, the use of the position information vector assists the proposed model while identifying those potential words that drew the focus of the whole sentence. According to the literature and the best of our knowledge, all prescribed deep features are utilized separately during aspect identification and their sentiment classification, whereas their joint combination is never analyzed. Therefore, the proposed approach is among those novel techniques that consider their collective utilization for accomplishing the aspect extraction and sentiment classification tasks. It can regard as the main novelty of the proposed approach. However, the use of the attention layer highlights the possible aspect and sentiment terms. The proposed method is evaluated through publicly available standard datasets for ABSA tasks comprising SemEval and Twitter. As a summary, this paper contributes the following:

1. The approach analyzes the combined utilization effect of word2vec embedding, contextual position information vector, and consistency-based tree structure on the deep learning-based algorithm performance while extracting aspects and categorizing their sentiments.
2. The proposed approach provides syntactic, logical, semantic, and contextual position information of sentences to the attention mechanism, which collectively enhances its ability to obtain the interaction among aspects, sentiment, and contextual words with the accurate placement of influenced terms.
3. The proposed approach utilizes the Attention-based Multi-Channel Convolutional Neural Network (MC-CNN) with various deep features to accurately identify aspects and classify their sentiments.
4. On standardized datasets including SemEval and Twitter, the proposed approach performs well based on evaluation metrics, such as precision, recall, and F1. Accordingly, the F1 measure shows 94% success in aspect term extraction and 92% in sentiment classification for the proposed approach.

The rest of the paper is structured as follows. Methodologies for extracting aspects and classifying sentiments are described as “Related Work” in [Section 2](#). “Proposed Research Methodology” describes the overall methodology and details of the proposed approach in [Section 3](#). According to [Section 4](#), the environment considered for developing this model is referred to as the “Experimental Arrangement”. This model’s performance comparison is presented in [Section 5](#) under “Results and Discussions”. In the final section, [Section 6](#), a conclusion is provided, and it describes the whole approach as a future direction for further research.

2 Related Work

In a state of technological developments of this era, the advancements in the field of social media (e.g., Facebook, LinkedIn, and Twitter) and E-commerce websites (e.g., Flipkart, Amazon, and eBay) allow their users to share their experiences regarding tour trips, movies, products, and hotels’ service in the form of opinion and reviews. On the other side, these expressions of feelings became the cause of a substantial volume of unstructured data, which is available on social networks and E-commerce websites, generated by their consumers or users. Analyzing such unstructured data unbiased is a crucial demand and challenge for individuals, organizations, and industries during their decision-making process [17,18]. There is a need for an automated system that mines the vast opinionated data for finalizing the decision criteria. In this regard, SA meets the desired need [19,20]. The main objective of SA is to acquire users’ attitudes, thoughts, and feelings towards individuals, groups, brands, and products [21]. Two well-known types of sentiment analysis are document-level sentiment analysis and sentence-level sentiment analysis. Additionally, document- and sentence-level SA cannot describe users’ dislikes or likes. Therefore, they only describe the overall sentiment expressed within that document and sentence. When a single sentence comprises more than one opinion or aspect, these two types did not meet the user’s requirements under these circumstances. It requires a particular type of SA, such as ABSA, that reviews aspects and their corresponding sentiment opinion available in a sentence. Therefore, ABSA provides more fine-grained details of the analysis. It describes users’ attitudes, likes, and dislikes regarding different aspects of a targeted entity [4,22]. On the other hand, these extracted attributes are of two types; they can be either explicit or implicit [16,23–26]. This contribution only considers explicit aspects, whereas the implicit attributes are beyond the scope.

The main subtasks of ABSA are known as aspect term extraction (ATE) and their sentiment classification (SC), which identify the aspects or opinionated targets from underlying textual reviews

and analyze their corresponding opinion sentiment. The aspect-oriented opinion expresses two kinds of knowledge. Consequently, the first one describes the sentiment polarity associated with the aspect. In the meantime, the second one justifies why the reviewer or user has associated such sentiment with this aspect. In addition, the earlier methodologies utilized these procedures for attributes and opinions extraction from online user-generated reviews [22]. Moreover, according to the literature, these aspects or features have been extracted through manually-based mechanisms. These are laborious, complicated, and time-consuming approaches that require a lot of effort [22,27–30]. In the early days of ABSA, machine learning methods, such as semi-supervised [31,32], supervised [33,34], and unsupervised [35], lexicons-based approaches (i.e., domain-based lexicons and SentiWordNets) [36–40], rule-based or pattern-based techniques [41–48], topic modelling based procedures [49–51], and tree or graph-based strategies [52,53] accomplished the feature and sentiment extraction task. Today, DL methods are famous for ABSA tasks, but including reasoning like the human brain remains an open research area.

Generally, two types of feature representations are famous in DL approaches for extracting aspects. These include the sequential structure and tree structure-based representations. The sequential expressions express the order of words in a sentence. Therefore, existing approaches depend on highly anticipating word2vec models for obtaining sequential indications. That depicts the words as low-dimension vectors and syntactic rules used to capture the semantic information, but they face issues regarding user-generated reviews [54]. Therefore, Ahmed et al. [55] proposed a method named, HEA-LSTM that identifies and extracts the potential aspects by combining hints with attention-based LSTM. Therefore, embedding indicators related to aspects' contextual-oriented facts make HEA-LSTM a highly focused approach. In the meantime, the contextual position and tree-based information can improve the performance of their technique. On the other side, existing methodologies have proven inadequate in useable and accurate identification and extraction of aspect terms. They rely on the textual context, which comprises contextual and long-term dependencies that cause to reduce their aspects' capturing abilities. In a state of these motivations, Chauhan et al. [56] presented a two-step hybrid unsupervised model, which combines an attention-based bi-LSTM model with linguistic patterns for accomplishing the task of aspect term extraction. In the meantime, the dependency and logical relations inclusion can improve their aspect extraction and identification accuracy. According to traditional approaches, aspect terms and contextual phrases are not appropriately considered in relation to their distance and the impact of interrelated contextual words and phrases. Hence, Huang et al. [14] presented a method named CPA-SA, which incorporates aspect-specific contextual location information while achieving ABSA tasks. Designed to reduce the inference between potential terms to determine their polarities on both sides, their function adjusts the weights of contextual words according to potential terms' position. This approach, however, excluded the influence of semantic and syntactic relationships during the accomplishment of this task.

In the related literature on ABSA, RNNs are noticed as inadequate in compatibilities regarding features compared to CNNs. These deficiencies motivate Da'u et al. [28] to proposed an approach comprised of MC-CNN that uses Part-of-Speech tags and word2vec embedding as textual features for targeting aspects extraction. However, although this approach performs well, they neglect the consideration of tree-based dependencies and contextual position information while accomplishing this task. Therefore, Xu et al. [57] proposed a novel approach named Dual Embedding CNN (DE-CNN), which comprises a CNN and two different embeddings for performing the task of aspect identification and extraction. One of these embeddings is pre-trained for general-purpose use, while the other is trained for a specific domain. While, traditionally, these aspects were extracted from user-generated reviews using manual annotations and rule creation approaches that made these methods

domain or task-dependent and highly challenging. In this situation, Barnaghi et al. [27] proposed a technique that automatically extracts these targeted aspects. Therefore, their approach utilizes a CNN with two embedding layers comprising GloVe. Between them, the first one is generally pre-trained, while the second one got trained according to a specific-domain perspective. They accomplish this task without considering any contextual position information and tree-based information. As a consequence of these deficiencies, the accuracy of their model can improve while predicting potential terms as aspects.

On the other hand, the tree-based representation has based upon the syntax-oriented, grammatical, and logical relations between the words' structure of sentences. Additionally, such a representation can better preserve the long-term dependencies relations between words compared to sequential feature representation. However, aspect extraction and their sentiment classification observe as being the utmost important task of ABSA. Consequently, various methods have contributed to the literature for accomplishing it. Among these methods, those that utilize the dependency tree structure express adequate performance. In addition, these methods consider only one-directional dependency tree propagation. These circumstances encouraged Luo et al. [58] to propose an approach that utilizes bidirectional dependency tree propagation for extracting aspects. Subsequently, this information is provided to Bi-LSTM and then forwarded to CRF, which uses sequential features with a tree-based structure for identifying aspect terms. However, consideration of contextual positional information and attention mechanisms can further improve model's accuracy. In another contribution, Wang et al. [59] proposed method encodes the information about propagation while ignoring the syntactical association between aspects and opinions words in a sentence. Eventually, dependency-based structural correspondence relations establishes for transferring cross-domain knowledge between source and target domains. In the meantime, an auto-encoder is used in their model for effective filtrating the noisy relation information. In addition, noticed limitations of RNN regarding tree-based dependencies handling motivate Ye et al. [10] to propose an approach named, DTBCSNN that uses its state-of-the-art capabilities for extracting aspects. The proposed CNN variant with linguistic characteristics is a DL-based algorithm and an end-to-end model that does not rely on manually-based characteristics, which extracts syntactic details from sentences using dependency parse trees. In a contribution, Lu et al. [5] proposed a method named, AGGCN that utilizes a graph-based CNN, which uses sentiment dependencies and syntactic information to encode aspect-specific gates that identify aspects and classify them as negative, positive, and neutral. Meanwhile, the utilization of word2vec embedding, the attention mechanism and positional information can enhance the accuracy of aspect identification.

Interpreting the above-expressed approaches, we concluded that all DL methodologies regarding aspects' identification and sentiment analysis have massively relied on either tree-based or word2vec embedding-based feature representations. Word2vec embedding is utilized for preserving the syntactical and semantical information of a targeted sentence. In the meanwhile, the tree-based structure uses to conserve grammatical and logical associations found among the words in a targeted sentence. Hence depending only on the word2vec embedding has become the cause of ignoring the grammatical and logical relations of the context. Meanwhile, depending upon the tree-based representation neglects the semantic and syntactic information of the targeted sentence. These circumstances motivated the proposed approach to developing the hypothesis that the technique that jointly considers the word2vec embedding along with the tree-based structure representations can better understand the syntactic, semantic, grammatical, and logical relation of a targeted sentence during the extraction of aspects and classifying their sentiments. In addition, concerning contextual positional information improves the proposed approach's ability to identify and extract targeted aspects' sentiment. In a state

of this inspiration, the proposed method collectively utilizes the bidirectional consistency tree (Bi-CTree) structure, word2vec embedding, and contextual positional information. These parameters are integrated and then passed to Att-MC-CNN, which learns the targeted aspects and their sentiment from a particular sentence. Eventually, these potential terms deliver to the Softmax function for predicting final attributes and their sentiment classification. The consideration of combined features differentiates it from the previous approach and makes it more comprehensive and robust than existing techniques.

3 Proposed Research Methodology

The existing methods rely separately upon the dependency tree structure and word embedding during aspect extraction and their sentiment classification. It becomes the prominent cause of considering a few contextual features such as syntactic and semantic information, whereas excluding other features such as grammatical and logical relation information. According to the relevant literature, the approach that considers all kinds of deep-features related to sentence structure performs better. In a state of such inspiration, Att-MC-CNN presented a methodology that considers bi-directional consistency tree (Bi-CTree) structure, word2vec embedding, and contextual positional information collectively as input. Att-MC-CNN accepts these parameters and learns the targeted aspects and their sentiments from a particular sentence. The use of the attention mechanism determines the emphasized terms, which highly influences the whole sentence context and then provides to the pooling layer. The extracted features pass towards Softmax for classifying potential terms as non-aspect or aspect terms identified from textual reviews. Depending on these identified aspect terms, Softmax determines their sentiment from context and classifies them as negative, positive, and neutral. This mechanism discriminates it from the previous procedures and makes it more comprehensive. The proposed methodology framework's detail has given in Fig. 1. During the last decades, the syntactic pattern techniques that utilize a dependency parser to generate linguistically and grammatical dependent rules have performed well in aspect identification and sentiment classification. In contrast, neither users nor websites impose prohibited constraints during the generation or sharing of feedback over invalid grammatical or syntactic patterns. It is, therefore, questionable whether these techniques can perform well when utilizing online user-generated data. However, consistency parsing techniques still perform well in these situations because consistency parsing depends upon the sentence-oriented syntactic structure. This tree structure still proves beneficial in mining those absolute parts that collectively formulate sentences even when the language's grammatical rules are violated. According to the literature, nouns or noun phrases are those categorical parts of context those primarily anticipated to be potential aspect terms. In addition, verbs, adverbs, and adjectives have expected as sentiment terms. These factors motivate the proposed approach to consider the consistency parser instead of the grammatical relational dependency parser. Due to this, the proposed method utilizes a bidirectional consistency tree to identify potential features from the targeted sentences. Thus, the entire sentence is divided into chunks. Among these chunks, the proposed approach considers only NP chunks, verbs, and adverbs, while it discards the remaining terms. These chunks eventually traverse as a tree, from right to left and left to right, for identifying and analyzing potential terms. The mathematical notation that expresses the whole task has given in the form of Eqs. (1)–(3).

$$ST_f = \text{Sigmoid} \left(\sum_{i=1}^l W_i x_i + b \right) \quad (1)$$

$$ST_b = \text{Sigmoid} \left(\sum_{i=1}^l W_i x_i + b \right) \quad (2)$$

$$Z = \text{Max}(ST_f, ST_b) \tag{3}$$

Here ST_f expresses a subtree that has traversed in the forward direction while ST_b subtree traversals in the backward direction. W_i is the corresponding weight of each term which dot product takes with the x_i . While b is the bias and l is the length of a subtree in Eqs. (1) and (2). These subtrees' output features identify with the help of *Sigmoid* functions. Then the finalized features are selected in the form of Z with the assistance of a *Max* function expressed in Eq. (3). It represents the selection of those features in which maximum influence observes among the subtree sequence in both forward or backward direction traversal of the tree.

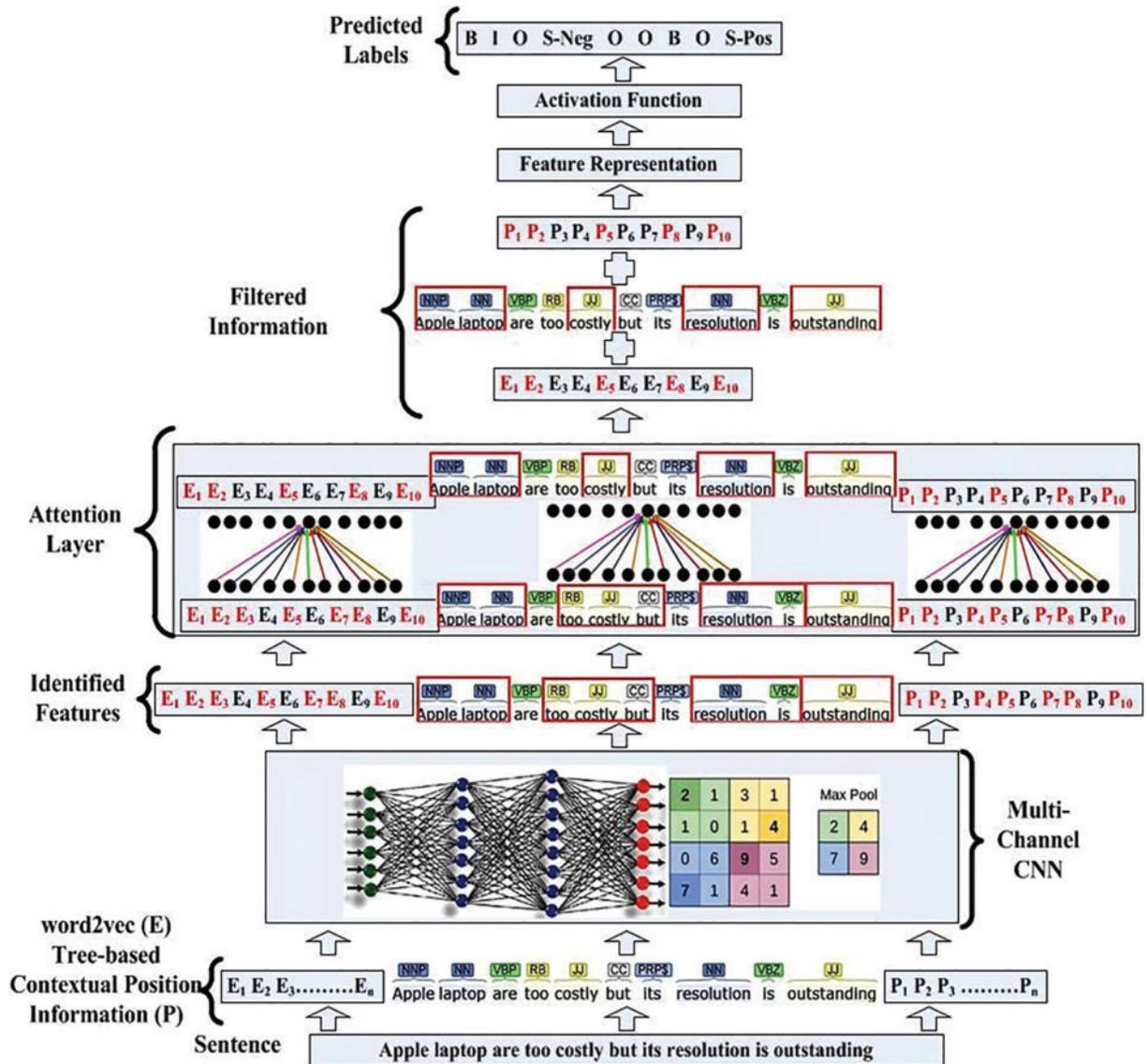


Figure 1: The attention-based MC-CNN's framework for ATE and SC

The sentence-oriented word position describes a critical factor in the NLP's tasks because it can express or change their degree of influence in the contextual information of a targeted sentence. Therefore, knowledge of the position-oriented information of words in a sentence has been considered significant during the model's training. Its reason is that when a similar word occurs at different positions in diverse contexts, it behaves differently in various places. In addition, it also produces diverse influences within context due to altering its position. Such as, when in a sentence "A" is previously noticed, it behaves like the word "B", when examining it next time, it acts like the word "C", which concludes that the position of any word expresses the diverse influence within the context of a sentence. Therefore, the proposed method uses the vector representation of contextual position information regarding each targeted sentence, of which mathematical notation is expressed in Eq. (4) [60] given below:

$$CPI_i = \lim_{j \rightarrow n} \left(\theta_j + (1 - \theta_j) \left(\frac{p_j}{n} \right) \right), 0 \leq \theta \leq 1, \quad (4)$$

In Eq. (4), CPI_i depicts the vector representation of contextual position information for i^{th} sentence. Meanwhile, the ratio θ_j reflects the j^{th} word's contribution within the range of $[0, 1]$ between position and presence. Furthermore, p_j depicts the word's spot in the focussed sentence, whereas the length of the sentence is denoted using n . The position-oriented influence of these factors is calculated using Eq. (4) if the j^{th} word exists in vocabulary. Alternatively, in case of absence, zero is added.

The understanding of human languages transforms into machines' insights through word embedding (i.e., representation of words). These low-level distributed n-dimensional representations have proven themselves a beneficial deep feature in solving various NLP's real-world scenarios. Although these representations can better capture or describe the syntactic and semantic relations between a couple of terms lying in a sentence. Therefore, using a special kind of relation-oriented vectors' offset has articulated these relationships. Consequently, word2vec embedding employs as an input parameter for Att-MC-CNN. In this way, pre-trained word2vec embedding on the Google News dataset, which contains vectors comprising three hundred dimensions, utilizes for this purpose. A padding mechanism is used to acquire a constant length of sentences. Thus achieving this goal zero-padding method is used, which maps each word with the corresponding embedding. Henceforth, a weight matrix W_{WE} is utilized during relations inferring process from word2vec embedding.

3.1 Multi-Channel Convolution Neural Network

The first channel of MC-CNN obtains selected features using Eq. (3) through the bi-directional consistency tree with the weight matrix W_T . Henceforth based on these features, MC-CNN learns the potential features using the Eq. (5) [61] given below:

$$C_T = \text{Leaky ReLU} (W_T x_T + b) \quad (5)$$

Here C_T depicts the features generated from the inclusion of tree-based representation, such as the consistency tree, *Leaky ReLU* is activation function, and W_T is the weight matrix. However, x_T represents the features extracted from Eq. (3), and b is the bias term. On the other hand, word2vec embedding is delivered on the MC-CNN's second channel with the weight matrix W_{WE} . Then, using the following Eq. (6), features are extracted.

$$C_{WE} = \text{Leaky ReLU} (W_{WE} x_{WE} + b) \quad (6)$$

The parameters of Eq. (6) are the same as represented in Eq. (5). Meanwhile, contextual position information vectors are provision upon the third channel of MC-CNN with weight matrix W_{CPI} .

Subsequently, Eq. (7) utilizes to identify the prominent features. This equation is parametrically similar to Eq. (5), which has given below:

$$C_{CPI} = \text{Leaky ReLU}(W_{CPI}x_{CPI} + b) \quad (7)$$

Furthermore, each channel comprises a convolutional and pooling layer that determines the potential features of the input. Max-pooling functions are used in the pooling layer to map features, and their maximum values indicate the attributes that are associated with them. Due to the fact that the highest value represents the most prominent terms, the highest value signifies the most important. This processing occurs in a parallel configuration.

3.2 Attention Layer

Additionally, different words have conceived diverse influence scores based upon the sentence's context. Therefore, the Att-MC-CNN has considered an attention layer to determine the highly effective terms of a sentence. The Att-MC-CNN supports the attention mechanism while providing syntactic, logical, semantic, and contextual position information of context, which collectively enhances its ability to obtain the interaction among aspects, sentiment, and contextual words with the accurate placement of influenced terms. Attention mechanisms generally consist of two modules. There are encoders and decoders. An encoder converts the targeted sentence into a real-valued equal length vector that provides semantic information for each term. In contrast, the decoder module is responsible for transforming encoded vectors and providing output. Attention mechanism's equations are expressed in Eqs. (8)–(10) [62] as given below:

$$u_p = \tanh(W_p C + b_p) \quad (8)$$

$$\alpha_p = \frac{\exp(u_p^T u_w)}{\sum_p \exp(u_p^T u_w)} \quad (9)$$

$$C_p = \sum_p \alpha_p C \quad (10)$$

In the equations above, u_w is expressing the context update vector. u_p represents the resultant of the hidden layer's vectors C , b_p , and W_p . Whereas W_p and b_p are the weight, and bias terms of the attention layer, respectively. In the last, α_p , represents the score of attention associated to each sentence word. After that, the Att-MC-CNN uses attention-mechanism to identify those terms that are highly prominent in each sentence. Att-MC-CNN learns regarding potential terms through attention-mechanism that identifies these terms in the light of tree-based structure, contextual position information, and word2vec embedding. The attention mechanism eventually filters out the most prominent terms in each sentence.

$$C_{Att-T} = \text{Attention}(C_T) \quad (11)$$

$$C_{Att-WE} = \text{Attention}(C_{WE}) \quad (12)$$

$$C_{Att-CPI} = \text{Attention}(C_{CPI}) \quad (13)$$

Here C_T , C_{WE} , and C_{CPI} are the outcomes of tree-based, word2vec-based, and contextual position information-based channels, respectively. In addition, C_{Att-T} , C_{Att-WE} , and $C_{Att-CPI}$ are the most prominent features of C_T , C_{WE} , and C_{CPI} , respectively filtered out based on attention mechanism.

3.3 Output Layer

Eventually, from each channel's outcome, filtered features along their context-based information are combined using a fusion layer and represented as feature representation which has been expressed mathematically in Eq. (14).

$$FR = C_{Att-T} \oplus C_{Att-WE} \oplus C_{Att-CPI} \quad (14)$$

$$Predicted\ aspects = softmax(W.FR + b) \quad (15)$$

Here C_{Att-T} , C_{Att-WE} , and $C_{Att-CPI}$ are the most prominent features of C_T , C_{WE} , and C_{CPI} , respectively. In addition, FR , \oplus , W , and b denote features' representation, combination function, term's weight and the bias term's representation, respectively. These resultants have delivered towards the Softmax function for classification and prediction purposes. Thus, all these terms classify as non-aspect or aspect terms. The architectural detail of the proposed approach has illustrated in Fig. 2. The manipulation of Eq. (15) emphasizes the predicted aspects of the targeted sentences and classifies them as non-aspects or aspects. As a result, the aspect-specific terms are admissible, while the non-aspect-specific terms are not. After that, Att-MC-CNN predicts the remaining terms' polarity as negative, positive, and neutral. Therefore, to determine their corresponding sentiments, these aspects are combined with their context. Eq. (16) illustrates this process. Thus, the Softmax function is used to integrate these predicted filtered terms with the results of Eq. (14) that classify their sentiments as negative, positive, and neutral. This procedure is observable from Eq. (17).

$$S = (Predicted\ aspects, FR) \quad (16)$$

$$Predicted\ sentiments = softmax(S) \quad (17)$$

Based on the whole section, we concluded that the proposed methodology uses plain text as its input. This plain text is then processed and analyzed to generate results and insights. As a consequence, this plain text is transformed into word2vec embedding vectors, tree structure-based vectors, and contextual positional information-based vectors. It is due to the fact that DL models cannot deal directly with plain text as input, and there is a need to transform or encode it into numbers or vectors. These vectors represent and organize data. In order to train and deploy our models, vectors and matrices represent inputs as numbers. In addition, proposed approach collectively utilize the word2vec embedding along with the tree-based structure representations to better understand the syntactic, semantic, grammatical, and logical relation of a targeted sentence. In the meantime, the inclusion of contextual positional information assists the proposed approach while to identify and extract targeted aspects' sentiment. Using this technique, the proposed approach gains a better understanding of each aspect and its different sentiments in the specific context. At the same time, the attention layer determines the context-impactful terms, which have the greatest impact in the context of the entire sentence, and delivers them along to the pooling layer. The pooling layer gathers these words and creates a representation of feature vectors. After these finalized features are generated, they are sent to Softmax for aspect-term identification. They are also classified as neutral, negative, or positive based on their context. Therefore, as an output, the proposed methodology provides aspect terms along with their classified sentiments as positive, negative, and neutral.

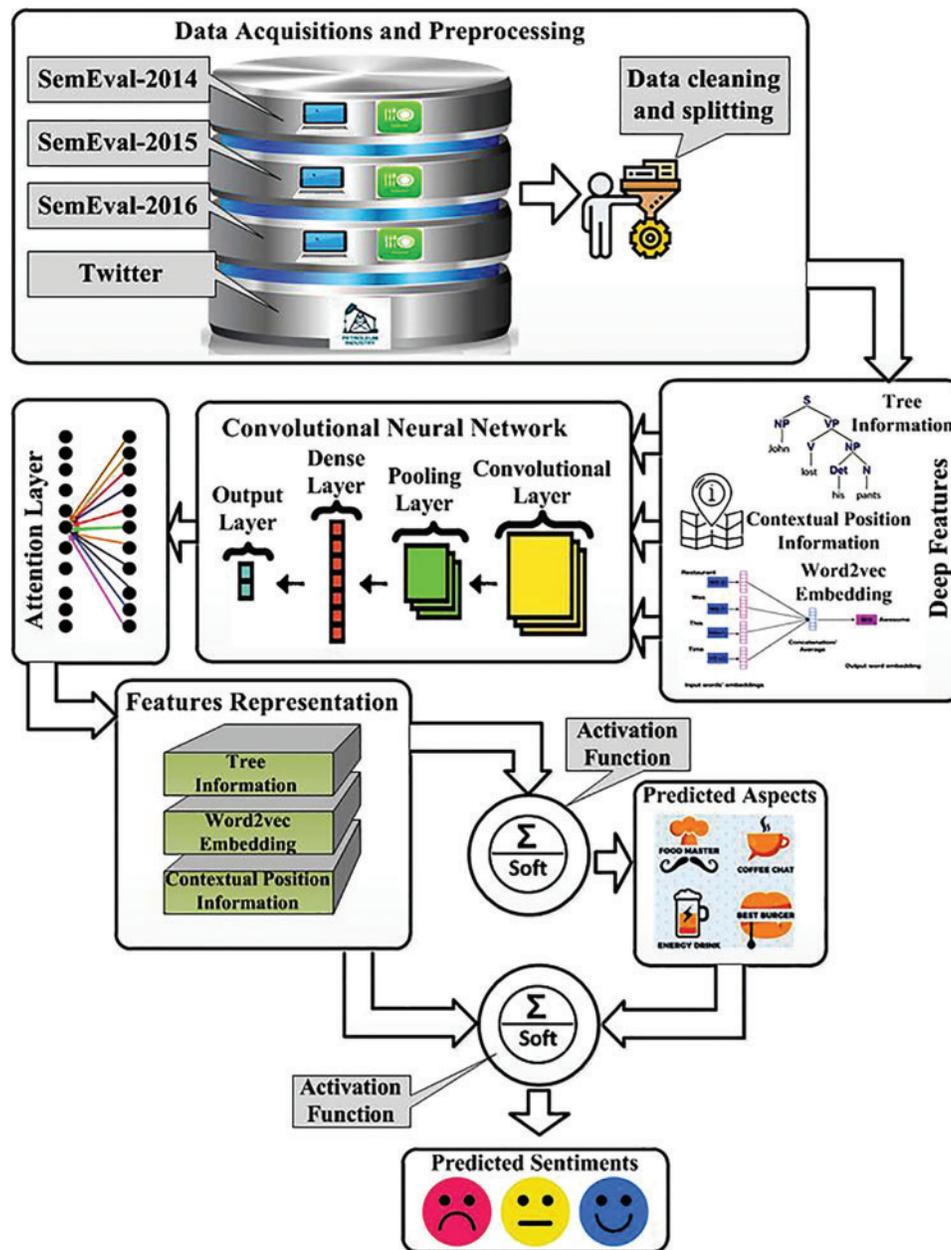


Figure 2: Proposed model of multi-channel convolutional neural network and attention mechanism

4 Experimental Arrangements

This paper briefly describes the setup of the environment for these experiments in this section. After that, the datasets used in these experiments for assessing the approach are discussed. Then pre-processing methods for filtering those datasets are discussed in detail, and lastly, the baselines utilized for comparison to that technique are expressed. Although precision, recall, and F1 measure metrics

are most frequently used in relevant literature of ABSA. Therefore, this study utilizes these metrics for evaluating this approach.

4.1 Environment for Experiments

To evaluate Att-MC-CNN performance, Windows 07 is utilized as the structured environment. The hardware consists of a GeForce GTX 1060 GPU, an Intel Xeon W3530 2.8 GHz CPU, and 16 GB of DDR3 RAM. Additionally, the Python-oriented GPU Tensorflow 2.0, Keras 2.1.0, and the PyCharm IDE for Python 3.7 tool are used in the implementation software.3.7.

4.2 Dataset Selection

On the basis of seven standard datasets, the proposed methodology's performance has been thoroughly assessed. SemEval organizers have made these test and training datasets available to the public. The laptop and restaurant domain reviews are found in the first two SemEval-14 datasets [63], respectively. The next two datasets, which are both part of the SemEval-15 [64] dataset, are related to restaurant and laptop.

The next two datasets comprised the laptop and restaurant domain which belongs to SemEval-16 [64]. The last one is the Twitter-Crude oil [65] dataset includes tweets regarding the investor's opinions about crude oil. This dataset has comprised of 1000 tweets. Yassine Hamdaoui produced this dataset, that's available since 2020. This study analyzes the proposed method performance on different domains, therefore, utilizes the crude oil domain to validate the efficiency of the proposed methodology. Table 1 shows the datasets and their statistics.

Table 1: Datasets and their statistics

Dataset	Train				Test			
	Aspect	Positive	Negative	Neutral	Aspect	Positive	Negative	Neutral
SemEval-14L	2358	1007	886	465	654	348	138	168
SemEval-14R	3693	2195	835	663	1134	742	196	196
SemEval-15L	1697	819	749	129	949	527	306	116
SemEval-15R	1192	536	503	153	678	354	246	78
SemEval-16L	2357	1150	749	458	914	661	209	44
SemEval-16R	1743	929	739	75	622	311	257	54
Twitter-crude oil	1789	839	749	201	623	311	264	48

4.3 Pre-Processing

Upon obtaining the textual reviews, the proposed method performed pre-processing to create a structured and clean textual dataset. Below is a description of the pre-processing:

- English sentences are converted to lower case words.
- Textual reviews' paragraphs are divided into sentences by the full-stop symbol.
- White spaces are used to divide sentences into tokens.
- All punctuation words are terminated from textual sentences.
- All words comprising an impure form of alphabetical (alphanumeric) characters have been removed.
- Removed all stop words from sentences.
- Removed all special characters from sentences.
- Terminated words with a length less than or equal to one character.

4.4 Baselines

The proposed methodology is compared with the baseline models. The details regarding the baseline models are presented in [Table 2](#), given below:

Table 2: An overview of the baseline methods considered for the proposed approach

Baseline methods/Reference	Models' novelty	ATE/SC	Publication year	Datasets
WDEmb [66]	Considering CRF along with a dependency-based approach.	ATE	2016	SemEval-14R, SemEval-15R and SemEval-14L
BiLSTM+CNN [67]	Character embedding has been encoded along with BiLSTM+CNN	ATE	2016	SemEval-14R, SemEval-15R and SemEval-14L
MTMN [68]	Multilayer attention-based model.	ATE	2017	SemEval-14R, SemEval-15R and SemEval-14L
MIN [69]	Consider two LSTMs along with the lexicon and dependency rules.	ATE	2018	SemEval-14R, SemEval-15R and SemEval-14L
SDAN [70]	Considering syntactic graph for extracting structural information along with CNN for aspect extraction.	ATE	2019	SemEval-14R, SemEval-15R and SemEval-14L
BiDTreeCRF [58]	Consider a tree-based approach along with Bi-LSTM and CRF.	ATE	2019	SemEval-14R, SemEval-15R and SemEval-14L
ESGCN [2]	They encode the graph edges related to syntactic dependencies and their interconnected labels.	ATE	2021	SemEval-14R, SemEval-15R and SemEval-14L
AGGCN [5]	Utilizes a graph-based CNN, dependencies and syntactic information.	SC	2021	SemEval-14R, 14L, SemEval-16R, and SemEval-15R

(Continued)

Table 2 (continued)

Baseline methods/Reference	Models' novelty	ATE/SC	Publication year	Datasets
GCN [6]	GCN uses graph-based CNN to infer structural and general semantics.	SC	2022	SemEval-14R, 14L, SemEval-16R, and SemEval-15R
Sentic GCN-BERT [7]	They merge SenticNet with a graph-based CNN.	SC	2022	SemEval-14R, 14L, SemEval-16R, and SemEval-15R
CPA-SAA [14]	Utilizes weights according to the position of word within context.	SC	2022	SemEval-14R, 14L, SemEval-16R, and SemEval-15R

5 Results and Discussions

The proposed methodology has noticed that the experiments' remarkable performance with regard to identification and extraction of aspects and their sentiment classification is the combined effect of the consistency tree structure information and word2vec embedding. The reason is that the consistency tree accurately determines those grammatical parts of the clause that, when combined, produce a sentence. It preserves the overall sequential information of targeted sentences and performs well even when violating the grammatical rules. In addition, word2vec embedding preserves sentence-oriented semantic and syntactic information. These factors concurrently are the reason for improving the efficiency of the Att-MC-CNN during experiments. Furthermore, the prediction of prominent terms within the context enhances through the utilization of contextual position information that supports these achievements. The combined effect of consistency tree, word2vec embedding, and contextual position information are represented through precision, recall, and F1-measure metrics in the consequent sections.

However, pre-training has been considered a significant factor in word2vec embedding's performance. Consequently, the best pre-training of word2vec embedding attains when being trained through massive datasets. Therefore, the proposed approach uses a pre-trained word2vec embedding using the dataset of Google news to obtain this objective. Alternatively, the Att-MC-CNN is evaluated using English-based datasets, such as the SemEval and Twitter datasets, which are respectively associated with laptops, restaurants, and crude oil. According to the proposed framework, the maximum sentence length is determined during training. Therefore, the maximum length of a single sentence in these reviews has been assumed to be that length. Zeros have been padded with all sentences whose conceived length is less than maximum length. Those procedures are taken into account by Att-MC-CNN to ensure that aspects and sentiments can be accurately identified. [Table 3](#) illustrates the detailed parameter setting of Att-MC-CNN.

Table 3: Setting of parameters for Att-MC-CNN

Proposed model parameters	Value
Word2vec embedding vector dimensions	300
Kernel sizes of CNN layer	(4,6,8)
Filter Sizes of CNN layer	(128,128,128)
Layers activation function	Leaky ReLU
Total epochs	50
Batch size	50
Name of optimizer	Adam
Learning rate	0.001
Loss function	Categorical cross entropy
Function for predictions	Softmax
Rate of dropout	0.5

The Att-MC-CNN thoroughly analyzed the arrangements of all the baseline approaches for conducting their experiments. Based on the analysis, the proposed method identifies which parameter settings improve or decrease performance. Therefore, relying on this parametric analysis and the technique of parameter optimization, the Att-MC-CNN deduces the best parameter settings from a thorough analysis of baseline approaches depicted in Table 2 and the parametric optimization technique. These parameter settings have illustrated in Table 3. Henceforth, a comparison of the proposed method to the baseline approaches based upon the F1 measure regarding aspect terms extraction is presented in Table 4, whereas their classification-based performance has demonstrated in Table 5. Based on its quantitative improvement across all domains of interest, Att-MC-CNN has achieved remarkable results when assessed against standardized datasets.

Table 4: Proposed approach F1 measure comparison to baseline approaches while extracting aspect terms

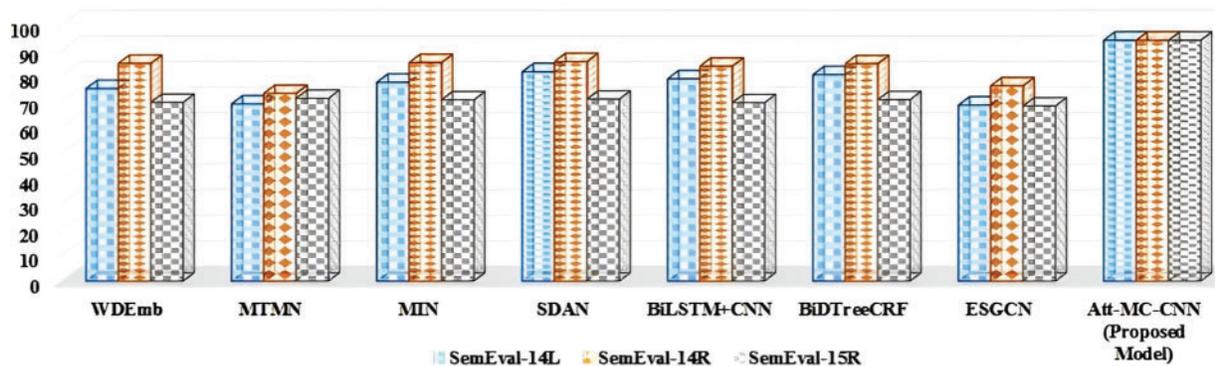
Model	SemEval-14L	SemEval-14R	SemEval-15R
WDEmb	75.16	84.97	69.73
MTMN	69.14	73.26	71.31
MIN	77.58	85.29	70.73
SDAN	81.68	85.68	71.01
BiLSTM+CNN	78.97	83.87	69.64
BiDTreeCRF	80.57	84.83	70.83
ESGCN	68.69	76.22	68.34
Att-MC-CNN (Proposed model)	94	94	94

Table 5: Proposed approach F1 measure comparison to baseline approaches while classifying sentiments

Model	SemEval-14L	SemEval-14R	SemEval-15R	SemEval-16R
CPA-SAF	70.87	72.81	60.26	71.47
CPA-SAA	71.5	73.38	60.15	72.43
AGGCN	68.99	73.82	64.51	73.92
Sentic GCN-BERT	79.05	81.03	71.28	79.56
GCN	74.63	77.35	66.39	75.43
BERT+ CRF	60.78	73.17	60.7	70.37
Att-MC-CNN (Proposed model)	84	87	92	90

5.1 Comparison of Performance during Aspect Term Extraction

The Att-MC-CNN expressed a remarkable performance increase in terms of F1 measure in the domain of SemEval-14L. The proposed model achieved reasonable margin success from all baseline approaches in the performance of the F1 score. Such as 25.31% from ESGCN, 13.43% from BiDTreeCRF, 15.03% from BiLSTM+CNN, 12.32% from SDAN, 16.42% from MIN, 24.86% from MTMN, and 18.84% from WDEmb. This performance gain can be observed in Fig. 3.

**Figure 3:** Comparison of baselines and proposed approach on different domains of dataset

The evaluation measures regarding the Att-MC-CNN have not comprised only one domain. Rather than this, the evaluation measure evaluates the Att-MC-CNN abilities through diverse domain datasets. Such procedure proves that the Att-MC-CNN is independent of the discipline, and the performance improvement does not restrict to a specific field. The consideration of combining inputs (contextual position information consistency tree, and word2vec embedding) becomes the reason of consistent performance against baseline approaches, even when the domain varies. Therefore, to determine the significant performance in the F1 measure, the Att-MC-CNN has been evaluated on the SemEval-14R's dataset too. The proposed model achieved reasonable margin success from all baseline approaches in the performance of the F1 score. Such as 17.78% from ESGCN, 9.17% from BiDTreeCRF, 10.13% from BiLSTM+CNN, 8.32% from SDAN, 8.71% from MIN, 20.74% from MTMN, and 9.03% from WDEmb. This remarkable performance is expressed in Fig. 3. In addition,

by vastly analyzing the performance of the Att-MC-CNN, the evaluation procedure also changes the dataset. Thus, the abilities of Att-MC-CNN has evaluated through the SemEval-15R dataset. The experimental results express a significant performance increase in F1-measure scores, such as 25.66% from ESGCN, 23.17% from BiDTreeCRF, 24.36% from BiLSTM+CNN, 22.99% from SDAN, 23.27% from MIN, 22.69% from MTMN, and 24.27% from WDEmb. The significant performance improvement has depicted in Fig. 3.

5.2 Comparison of Performance during Sentiment Classification

In the SemEval-14L dataset, the proposed model performed significantly better in the F1 measure score, which is 84%. In terms of F1 measure performance, it achieves an appropriate success in margin over all baseline approaches. Specifically, their performance enhancement is 23.22%, 9.37%, 4.95%, 15.01%, 12.5%, and 13.13% compared to BERT+CRF, GCN, Sentic GCN-BERT, AGGCN, CPA-SAA, and CPA-SAF, respectively. Fig. 4 shows these performance improvements.

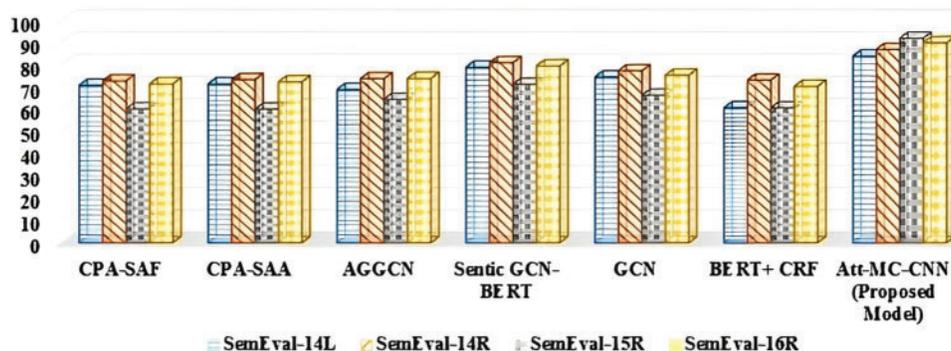


Figure 4: Comparison of baselines and proposed approach on different domains of dataset

The Att-MC-CNN obtained an F1 measure score of 87% when identified aspects from the SemEval-14R dataset are classified. Therefore, it achieves an improvement, which is 13.83, 9.65, 5.97, 13.18, 13.62, and 14.19 compared to BERT+CRF, GCN, Sentic GCN-BERT, AGGCN, CPA-SAA, and CPA-SAF, respectively. A pictorial representation of these performance enhancements is shown in Fig. 4. Att-MC-CNN also performs better than prescribed baselines in the SemEval-15R dataset, which is 92%. Accordingly, the performance enhancement obtained from baselines BERT+CRF, GCN, Sentic GCN-BERT, AGGCN, CPA-SAA, and CPA-SAF is 31.3, 25.61, 20.72, 27.49, 31.85, and 31.74, respectively. In relation to all considered baselines, the proposed approach performs more reasonably, which can be seen in Fig. 4. On the SemEval-16R dataset, however, the Att-MC-CNN achieves an improvement of 90% in F1. Consequently, it improves 19.63%, 14.57%, 10.44%, 16.08%, 17.57%, and 18.53% compared to BERT+CRF, GCN, Sentic GCN-BERT, AGGCN, CPA-SAA, and CPA-SAF, respectively. The collective utilization of tree-based and word2vec embedding along with contextual position information enhances the attention mechanism's precision while identifying relations among aspect, sentiment, and their corresponding context. As a result of the collective use of deep features, the proposed approach is able to make more accurate predictions. This efficiency-based improvement is illustrated graphically in Fig. 4.

5.3 Other Metrics and Performance of Proposed Approach

The Att-MC-CNN's ATE and SC tasks' performance improvement is estimated using a variety of metrics and dataset domains. The Att-MC-CNN's performance is evaluated through precision

metrics based on previously discussed datasets that indicate significant performance improvements. Accordingly, ATE's task achieved 95% precision over three domains of dataset, namely SemEval-15R, SemEval-15L, and SemEval-14L. Its minimum precision is 90% on SemEval-14R's domain. Meanwhile, the SC task achieved maximum precision, which is 91% on SemEval-15R's. Meanwhile, its minimum precision value is 80% on the SemEval-14L domain dataset. Fig. 5 shows precision-related observations of Att-MC-CNN using other dataset domains.

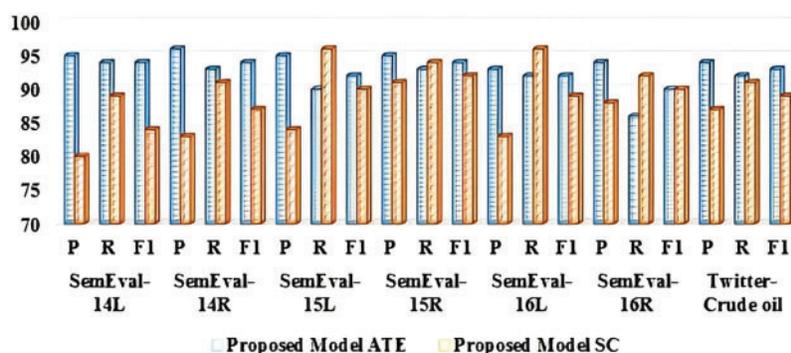


Figure 5: Model F1 measure precision, and recall based performance upon different datasets

However, the experiments analyze the Att-MC-CNN's tasks performance improvement using the recall metric through variant datasets comprising diverse domains. Therefore, the maximum recall of ATE task determined on SemEval-14L's is 94%. Meanwhile, its minimum recall value analyzed on SemEval-16R's is 86%. Besides, SC achieved a maximum recall value of 96% on SemEval-15L and 16L's dataset's domain. While the minimum recall value for SemEval-14L's is 84%. In Fig. 5, we demonstrate recall-based observations of Att-MC-CNN for other dataset domains. A variety of metrics are included in the experimental analysis, such as the F1 measure, precision, and recall. These metrics determine how well Att-MC-CNN performs on different datasets and domains. Thus, Att-MC-CNN's assessed capabilities through precision and recall metrics on diverse datasets demonstrate significant improvements in performance. Therefore, the maximum F1 value of ATE task analysis on three datasets, comprising SemEval-14R, 14L, and 15R's domains, is 94%. Meanwhile, the F1 minimum score found in the domain of SemEval-16R is 90%. On SemEval-15R's domain dataset, SC's task received the highest F1 score of 92%. According to the domain dataset of SemEval-14L, it has a minimum F1 score of 84%. Fig. 5 shows the F1 measure observed during experiments regarding Att-MC-CNN on other datasets.

The Att-MC-CNN illustrates a novel methodology that learns aspects implicitly rather than explicitly as input. By identifying these aspects, the proposed method predicts their sentiments based on relevant contexts. Att-MC-CNN's performance and accuracy of sentiment classification are evaluated based on a variety of metrics, such as precision, recall, and F1 measures. When performing the ATE task, the proposed approach produces a better precision value, while the SC task yields a better recall value. It shows mixed achievement trends for ATE and SC tasks in the F1 measure. Compared to baseline approaches, the framework achieves significant improvements in all dataset fields. This approach has shown better performance because word2vec embeddings, consistency trees, and contextual position information are considered as inputs for MC-CNN. At the same time, attention mechanisms enhance sentiment classification accuracy and targeted aspects identification. Using deep features in combination, these quantitative achievements demonstrate the importance of combining deep features across various domains.

6 Conclusion

In this approach, Attention-based MC-CNN is used to identify the targeted aspect and classify sentiments based on textual reviews. The consideration of word2vec embedding, contextual position information consistency, and of utilization tree-based structure becomes the cause of the significant performance of the proposed approach. Moreover, combining this information into a single framework helps to capture rich and complex semantic associations between words, leading to a more accurate sentiment classification. Models that incorporate such information are more robust than those that use a single feature. Furthermore, models that incorporate such information are less prone to overfitting and can more accurately generalize to unseen data. Consequently, the evaluation depicted through experimental results supports the hypothesis that the combined utilization of word2vec embedding, tree-based structure, and contextual position information vectors improves the accuracy of tasks such as aspect extraction and sentiment classification. The collective utilization of deep features improves the attention mechanism's ability to determine the relationships among aspects, sentiments, and their corresponding context. This research will be extended in future work by determining the implicit polarity related to specific aspects.

Funding Statement: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. 3418].

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] J. Yang, R. Yang, H. Lu, C. Wang and J. Xie, "Multi-entity aspect-based sentiment analysis with context, entity, aspect memory and dependency information," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 18, no. 4, pp. 1–22, 2019.
- [2] S. Wu, H. Fei, Y. Ren, B. Li, F. Li *et al.*, "High-order pair-wise aspect and opinion terms extraction with edge-enhanced syntactic graph convolution," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2396–2406, 2021.
- [3] Y. Ma, H. Peng, T. Khan, E. Cambria and A. Hussain, "Sentic LSTM: A hybrid network for targeted aspect-based sentiment analysis," *Cognitive Computation*, vol. 10, no. 4, pp. 639–650, 2018.
- [4] P. Karagoz, B. Kama, M. Ozturk, I. H. Toroslu and D. Canturk, "A framework for aspect based sentiment analysis on Turkish informal texts," *Journal of Intelligent Information Systems*, vol. 53, no. 3, pp. 431–451, 2019.
- [5] Q. Lu, Z. Zhu, G. Zhang, S. Kang and P. Liu, "Aspect-gated graph convolutional networks for aspect-based sentiment analysis," *Applied Intelligence*, vol. 51, no. 7, pp. 4408–4419, 2021.
- [6] A. Dai, X. Hu, J. Nie and J. Chen, "Learning from word semantics to sentence syntax by graph convolutional networks for aspect-based sentiment analysis," *International Journal of Data Science and Analytics*, vol. 14, no. 1, pp. 1–10, 2022.
- [7] B. Liang, H. Su, L. Gui, E. Cambria and R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks," *Knowledge-Based Systems*, vol. 235, no. 4, pp. 107643, 2022.
- [8] C. Gan, L. Wang, Z. Zhang and Z. Wang, "Sparse attention based separable dilated convolutional neural network for targeted sentiment analysis," *Knowledge-Based Systems*, vol. 188, pp. 104827, 2020.
- [9] M. Ramzan, H. U. Khan, S. M. Awan, W. Akhtar, M. Ilyas *et al.*, "A survey on using neural network based algorithms for hand written digit recognition," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 9, pp. 519–528, 2018.

- [10] H. Ye, Z. Yan, Z. Luo and W. Chao, "Dependency-tree based convolutional neural networks for aspect term extraction," in *Proc. of Pacific-Asia Conf. on Knowledge Discovery and Data Mining*, Jeju, South Korea, pp. 350–362, 2017.
- [11] M. Giménez, J. Palanca and V. Botti, "Semantic-based padding in convolutional neural networks for improving the performance in natural language processing. A case of study in sentiment analysis," *Neurocomputing*, vol. 378, no. 7, pp. 315–323, 2020.
- [12] S. Yu, D. Liu, W. Zhu, Y. Zhang and S. Zhao, "Attention-based LSTM, GRU and CNN for short text classification," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 1, pp. 1–8, 2020.
- [13] J. Zhang, F. Liu, W. Xu and H. Yu, "Feature fusion text classification model combining CNN and BiGRU with multi-attention mechanism," *Future Internet*, vol. 11, no. 11, pp. 237, 2019.
- [14] B. Huang, R. Guo, Y. Zhu, Z. Fang, G. Zeng *et al.*, "Aspect-level sentiment analysis with aspect-specific context position information," *Knowledge-Based Systems*, vol. 243, no. 2, pp. 108473, 2022.
- [15] W. Meng, Y. Wei, P. Liu, Z. Zhu and H. Yin, "Aspect based sentiment analysis with feature enhanced attention CNN-BiLSTM," *IEEE Access*, vol. 7, pp. 167240–167249, 2019.
- [16] C. Yang, H. Zhang, B. Jiang and K. Li, "Aspect-based sentiment analysis with alternating coattention networks," *Information Processing & Management*, vol. 56, no. 3, pp. 463–478, 2019.
- [17] V. Gupta, V. K. Singh, P. Mukhija and U. Ghose, "Aspect-based sentiment analysis of mobile reviews," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 4721–4730, 2019.
- [18] A. Mahmood, H. U. Khan and M. Ramzan, "On modelling for bias-aware sentiment analysis and its impact in Twitter," *Journal of Web Engineering*, vol. 19, no. 1, pp. 1–28, 2020.
- [19] K. Vivekanandan and J. S. Aravindan, "Aspect-based opinion mining: A survey," *International Journal of Computer Applications*, vol. 106, no. 3, pp. 975–980, 2014.
- [20] H. U. Khan, S. Nasir, K. Nasim, D. Shabbir and A. Mahmood, "Twitter trends: A ranking algorithm analysis on real time data," *Expert Systems with Applications*, vol. 164, no. 3, pp. 113990, 2021.
- [21] F. Tang, L. Fu, B. Yao and W. Xu, "Aspect based fine-grained sentiment analysis for online reviews," *Information Sciences*, vol. 488, no. 6, pp. 190–204, 2019.
- [22] M. Al-Smadi, B. Talafha, M. Al-Ayyoub and Y. Jararweh, "Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2163–2175, 2019.
- [23] V. Gupta, V. K. Singh, P. Mukhija and U. Ghose, "Aspect-based sentiment analysis of mobile reviews," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 1–10, 2019.
- [24] M. Tubishat, N. Idris and M. A. Abushariah, "Implicit aspect extraction in sentiment analysis: Review, taxonomy, opportunities, and open challenges," *Information Processing & Management*, vol. 54, no. 4, pp. 545–563, 2018.
- [25] M. Al-Smadi, M. Al-Ayyoub, Y. Jararweh and O. Qawasmeh, "Enhancing aspect-based sentiment analysis of Arabic hotels' reviews using morphological, syntactic and semantic features," *Information Processing & Management*, vol. 56, no. 2, pp. 308–319, 2019.
- [26] X. Ma, J. Zeng, L. Peng, G. Fortino and Y. Zhang, "Modeling multi-aspects within one opinionated sentence simultaneously for aspect-level sentiment analysis," *Future Generation Computer Systems*, vol. 93, no. 1–2, pp. 304–311, 2019.
- [27] P. Barnaghi, G. Kontonatsios, N. Bessis and Y. Korkontzelos, "Aspect extraction from reviews using convolutional neural networks and embeddings," in *Proc. of Int. Conf. on Applications of Natural Language to Information Systems*, Salford, UK, pp. 409–415, 2019.
- [28] A. Da'u and N. Salim, "Aspect extraction on user textual reviews using multi-channel convolutional neural network," *PeerJ Computer Science*, vol. 5, no. 6, pp. e191, 2019.
- [29] N. Liu, B. Shen, Z. Zhang, Z. Zhang and K. Mi, "Attention-based sentiment reasoner for aspect-based sentiment analysis," *Human-Centric Computing and Information Sciences*, vol. 9, no. 1, pp. 35, 2019.
- [30] D. Zeng, Y. Dai, F. Li, J. Wang and A. K. Sangaiah, "Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 3971–3980, 2019.

- [31] Z. Yan, M. Xing, D. Zhang and B. Ma, "EXPRS: An extended pagerank method for product feature extraction from online consumer reviews," *Information & Management*, vol. 52, no. 7, pp. 850–858, 2015.
- [32] A. K. Samha, Y. Li and J. Zhang, "Aspect-based opinion extraction from customer reviews," *Computer Science & Information Technology*, vol. 4, pp. 149–160, 2014.
- [33] N. Jihan, Y. Senarath, D. Tennekoon, M. Wickramaratne and S. Ranathunga, "Multi-domain aspect extraction using support vector machines," in *Proc. of the 29th Conf. on Computational Linguistics and Speech Processing (ROCLING)*, Taipei, Taiwan, pp. 308–322, 2017.
- [34] R. Hegde and S. Seema, "Aspect based feature extraction and sentiment classification of review data sets using Incremental machine learning algorithm," in *Proc. of 2017 Third Int. Conf. on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, Chennai, India, pp. 122–125, 2017.
- [35] Y. Xiang, H. He and J. Zheng, "Aspect term extraction based on MFE-CRF," *Information*, vol. 9, no. 8, pp. 198, 2018.
- [36] C. Liao, C. Feng, S. Yang and H. Y. Huang, "A hybrid method of domain lexicon construction for opinion targets extraction using syntax and semantics," *Journal of Computer Science and Technology*, vol. 31, no. 3, pp. 595–603, 2016.
- [37] M. E. Mowlaei, M. S. Abadeh and H. Keshavarz, "Aspect-based sentiment analysis using adaptive aspect-based lexicons," *Expert Systems with Applications*, vol. 148, no. C–1, pp. 113234, 2020.
- [38] M. S. Wai and S. S. Aung, "Simultaneous opinion lexicon expansion and product feature extraction," in *Proc. of IEEE/ACIS 16th Int. Conf. on Computer and Information Science (ICIS)*, Wuhan, China, pp. 107–112, 2017.
- [39] X. Fu, L. Guo, Y. Guo and Z. Wang, "Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon," *Knowledge-Based Systems*, vol. 37, no. 88–94, pp. 186–195, 2013.
- [40] M. L. Yadav and B. Roychoudhury, "Effectiveness of domain-based lexicons vis-à-vis general lexicon for aspect-level sentiment analysis: A comparative analysis," *Journal of Information & Knowledge Management*, vol. 18, no. 3, pp. 1950033, 2019.
- [41] Q. Liu, Z. Gao, B. Liu and Y. Zhang, "Automated rule selection for opinion target extraction," *Knowledge-Based Systems*, vol. 104, no. 1, pp. 74–88, 2016.
- [42] Y. Kang and L. Zhou, "RubE: Rule-based methods for extracting product features from online consumer reviews," *Information & Management*, vol. 54, no. 2, pp. 166–176, 2017.
- [43] M. Z. Asghar, A. Khan, S. R. Zahra, S. Ahmad and F. M. Kundi, "Aspect-based opinion mining framework using heuristic patterns," *Cluster Computing*, vol. 22, no. S3, pp. 7181–7199, 2019.
- [44] A. Kushwaha and S. Chaudhary, "Review highlights: Opinion mining on reviews. A hybrid model for rule selection in aspect extraction," in *Proc. of the 1st Int. Conf. on Internet of Things and Machine Learning*, Liverpool, United Kingdom, pp. 1–6, 2017.
- [45] Q. Liu, Z. Gao, B. Liu and Y. Zhang, "Automated rule selection for aspect extraction in opinion mining," in *Proc. of Twenty-Fourth Int. Joint Conf. on Artificial Intelligence*, Buenos Aires Argentina, pp. 1291–1297, 2015.
- [46] T. A. Rana and Y. N. Cheah, "Hybrid rule-based approach for aspect extraction and categorization from customer reviews," in *2015 9th Int. Conf. on IT in Asia (CITA)*, Sarawak, Malaysia, pp. 1–5, 2015.
- [47] T. A. Rana and Y. N. Cheah, "A two-fold rule-based model for aspect extraction," *Expert Systems with Applications*, vol. 89, no. 5, pp. 273–285, 2017.
- [48] F. Z. Ruskanda, D. H. Widiantoro and A. Purwarianti, "Comparative study on language rule based methods for aspect extraction in sentiment analysis," in *Proc. of 2018 Int. Conf. on Asian Language Processing (IALP)*, Bandung, Indonesia, pp. 56–61, 2018.
- [49] M. Shams and A. Baraani-Dastjerdi, "Enriched LDA (ELDA): Combination of latent Dirichlet allocation with word co-occurrence analysis for aspect extraction," *Expert Systems with Applications*, vol. 80, no. 4, pp. 136–146, 2017.

- [50] S. J. Das and B. Chakraborty, "An approach for automatic aspect extraction by Latent Dirichlet Allocation," in *Proc. of IEEE 10th Int. Conf. on Awareness Science and Technology (iCAST)*, Morioka, Japan, pp. 1–6, 2019.
- [51] C. Wan, Y. Peng, K. Xiao, X. Liu, T. Jiang *et al.*, "An association-constrained LDA model for joint extraction of product aspects and opinions," *Information Sciences*, vol. 519, no. 3, pp. 243–259, 2020.
- [52] G. Ansari, C. Saxena, T. Ahmad and M. Doja, "Aspect term extraction using graph-based semi-supervised learning," *Procedia Computer Science*, vol. 167, pp. 2080–2090, 2020.
- [53] S. W. Jeon, H. J. Lee, H. Lee and S. Cho, "Graph based aspect extraction and rating classification of customer review data," in *Proc. of Int. Conf. on Database Systems for Advanced Applications*, Chiang Mai, Thailand, pp. 186–199, 2019.
- [54] S. Rida-E-Fatima, A. Javed, A. Banjar, A. Irtaza, H. Dawood *et al.*, "A multi-layer dual attention deep learning model with refined word embeddings for aspect-based sentiment analysis," *IEEE Access*, vol. 7, pp. 114795–114807, 2019.
- [55] M. Ahmed, Q. Chen, Y. Wang and Z. Li, "Hint-embedding attention-based LSTM for aspect identification sentiment analysis," in *Proc. of Pacific Rim Int. Conf. on Artificial Intelligence*, Cuvu, Yanuca Island, Fiji, pp. 569–581, 2019.
- [56] G. S. Chauhan, Y. K. Meena, D. Gopalani and R. Nahta, "A two-step hybrid unsupervised model with attention mechanism for aspect extraction," *Expert Systems with Applications*, vol. 161, no. 1–2, pp. 113673, 2020.
- [57] H. Xu, B. Liu, L. Shu and P. S. Yu, "Double embeddings and cnn-based sequence labeling for aspect extraction," in *Proc. of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia, pp. 592–598, 2018.
- [58] H. Luo, T. Li, B. Liu, B. Wang and H. Unger, "Improving aspect term extraction with bidirectional dependency tree representation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 7, pp. 1201–1212, 2019.
- [59] W. Wang and S. J. Pan, "Syntactically meaningful and transferable recursive neural networks for aspect and opinion extraction," *Computational Linguistics*, vol. 45, no. 4, pp. 705–736, 2020.
- [60] Y. Jin, H. Zhang and D. Du, "Incorporating positional information into deep belief networks for sentiment classification," in *Proc. of Industrial Conf. on Data Mining*, New York, NY, USA, pp. 1–15, 2017.
- [61] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The Handbook of Brain Theory and Neural Networks*, vol. 3361, no. 10, 1995.
- [62] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola *et al.*, "Hierarchical attention networks for document classification," in *Proc. of the Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, California, pp. 1480–1489, 2016.
- [63] D. Kirange, R. R. Deshmukh and M. Kirange, "Aspect based sentiment analysis semeval-2014 task 4," *Asian Journal of Computer Science and Information Technology*, vol. 4, pp. 72–75, 2014.
- [64] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar and I. Androutsopoulos, "Semeval-2015 task 12: Aspect based sentiment analysis," in *Proc. of the 9th Int. Workshop on Semantic Evaluation (SemEval)*, Denver, Colorado, pp. 486–495, 2015.
- [65] Z. Jiang, L. Zhang, L. Zhang and B. Wen, "Investor sentiment and machine learning: Predicting the price of China's crude oil futures market," *Energy*, vol. 247, no. 2, pp. 123471, 2022.
- [66] Y. Yin, F. Wei, L. Dong, K. Xu, M. Zhang *et al.*, "Unsupervised word and dependency path embeddings for aspect term extraction," in *Proc. of IJCAI Int. Joint Conf. on Artificial Intelligence*, New York, USA, vol. 2016, pp. 2979–2985, 2016.
- [67] X. Ma and E. Hovy, "End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF," in *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics*, Berlin, Germany, pp. 1064–1074, 2016.
- [68] W. Wang, S. J. Pan and D. Dahlmeier, "Multi-task memory networks for category-specific aspect and opinion terms co-extraction," in *Proc. of 31st Conf. on Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, pp. 3316–3322, 2017.

- [69] X. Li and W. Lam, “Deep multi-task learning for aspect term extraction with memory interaction,” in *Proc. of the Conf. on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, pp. 2886–2892, 2017.
- [70] J. Zhang, G. Xu, X. Wang, X. Sun and T. Huang, “Syntax-aware representation for aspect term extraction,” in *Proc. of Pacific-Asia Conf. on Knowledge Discovery and Data Mining*, Macau, China, pp. 123–134, 2019.